Data Warehousing Concept

Name

Institutional Affiliation

Data Warehousing Concept

**Introduction**

In computing, data warehouse is a system that is purposely used for reporting as well as analysis, which is known to be a core component of business intelligence. Ideally, data warehouses are core repositories of integrated data from one or more disparate sources. Historical and current data is stored in a single place within the data warehouse and are used to for creating analytics report for customers and workers throughout the enterprise. Notably, the data from the warehouse is stored in the data warehouse and then uploaded from the operational systems such as sales and marketing. In ensuring data cleaning, data before being stored in the warehouse must pass operational systems. This is critical is critical for data operations in ensuring that data is of high quality before it is used in the data warehouse for reporting. Data warehouse comprises of data integrated sources for preparing different forms of analytical reports which is then distributed to the knowledge workers within the enterprise. This paper aims to discuss the concept of data warehouse within the domain of technology and associated business rules as well as a complete schema diagram together with the justification.

In the technology domain, data warehouse that operates on a typical Extract, Transform, Load (ETL) approach, utilizes staging applications and databases, access, and integration layers to undertake their respective functions. Additionally, with the staging databases they keep raw data emanating from each data source then the integrating layers integrate it. Similarly, the technology domain facilitates integration of data to ensure that it is further arranged based on a hierarchical structure also known as dimensions. In the same way, the data that is catalogued is made readily available to the managers and systems users who have such rights and privileges for data access. This is paramount in enabling such managers and professionals to perform various activities such as data mining activities, decision support, as well as other operations such as market research.

 Database warehouse is considered to be the future of every company from the technological point of view. Such system combines methodologies, user management system, data manipulation and technologies used to generate insights relating to the company. Additionally, it is considered to be repositories of data from many sources which keeps not only current but also historical data. This data create analytical reports both annually or quarterly in nature depending on the needs of the business. Regarding features that are provided by data warehouses, they including providing organizations with a comprehensive decision support which has ideal reporting features for the organizations. Another future provided is that it is helpful in subject orientation, whereby data is obtained from different sources and a company can use specific data applicable to its own field. This is critical in helping the companies to gain more insights on how data can be applied in a manner that is beneficial to all sectors within the company.

**Business Rules**

Business rules and data warehouses are key concepts and technologies that have implications on a wide varieties of company’s tasks. From the technological domain within the data warehouse, application of the business mining approach to system sourcing as well as data analytical concept is key in addressing the gaps in data analysis and can accelerate the analysis and reliability of data. The table below represents business rules in the technology domain and in line with the data warehouse development lifecycle, whereby business rules are applicable.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Business Rule Mining** | **Data Analysis/ Data Quality** | **Logical Data Modelling** |
| Terms (data elements) |  |  X | X |
| Facts (Data relationships |  | X | X |
| Constraints (Data elements values) | X | X | X |
| Inferred knowledge (Understanding something about data from another data element or state) | X |  |  |
| Computational Rules | X |  |  |
| Action enabling | X |  |  |

**Snowflake Scheme and ERD**

This section capture ERD schema for the chosen domain. In this schema, it capture data from different sources that is mostly revolving around customer, product, sales, and data storage but from the technological point of view.

 Merging the business rules together with the star schema is configured and enabled through a concept known as star schema. Ideally, this merging occurs as follows:

1. When each dimension is a constraint in the business rule.
2. When all facts are constraint in the business rules.
3. When all actions refer to the combination of facts and combinations of facts as well as dimensions relating to a star schema instance.

Ideally, a typical rule concept that relates to star schema instances can be expresses as shown below:

 IF constraint\_1 operator value AND

 Constraint\_2 operator value AND….

 THEN action

IF act as a representative of the attributes within the each table that are participating in the star schema and that are needed in validation for the desired action. In the constraint section, constraint\_1 and constraint\_2 depicts dimension and the fact table attributes, while the operator act as the conditional operator.

References

Gour, V., Sarangdevot, S. S., Tanwar, G. S., & Sharma, A. (2010). Improve performance of extract, transform and load (ETL) in data warehouse. *International Journal on Computer Science and Engineering*, *2*(3), 786-789.

Kaula, R. (2009). Business rules for data warehouse. *International Journal of Information Technology*, *5*(1), 58-66.